BIM BASED CLASSIFICATION OF BUILDING PERFORMANCE DATA FOR ADVANCED ANALYSIS

Stephan Hoerster[1]; Karsten Menzel[1]

1: IRUSE Department, Western Gateway Building, University College Cork, Ireland

ABSTRACT

Data density has increased dramatically through utilization of wireless sensors and smart metering devices. While smart meters are commonly monitored in 15 minute intervals, wireless devices are usually polled at much higher rates. Due to these frequent reading intervals, the high data granularity creates a need for improved analysis and categorisation of data.

This research aims to extract descriptive information from BIM (Building Information Modelling) models in order to classify and structure monitored data. This is achieved through application of the open BIM standard IFC (Industry Foundation Classes). The extracted information gets combined with current DWH (Data Warehouse) methodologies to support superior and efficient analysis. The DWH model is in sync with the IFC standard to enable a straight forward mapping of IFC objects into the DWH.

Several KPIs (Key Performance Indicators) are be applied that support the combination of monitored fact data with the extracted descriptive information. The analysis of monitored data is realised through DWH cubes. These cubes are modelled with selected IFC objects in order to maximise their significance. The envisaged DWH allows the categorisation of data per (i) building system, (ii) building tenant, (iii) building space and (iv) time. Moreover, the DWH allows any combination of afore mentioned categories based on various time intervals. This enables various analysis scenarios, e.g. the calculation of the average electricity consumption per tenant or the peak temperature in a building level.

The benefit of reusing selected objects from BIM models over existing solutions is the elimination of repetitive work. In the past, descriptive data was acquired multiple times. This work demonstrates that data extracted from BIM models can be successfully used to enrich data analysis in a DWH system. The major achievement of this work is the elimination of the need to populate static information in the DWH through other means.

This concept is evaluated with data sets obtained from real buildings. A proof-of-concept DWH is presented in this work. The main obstacles during the ETL (Extract, Transform and Load) process of individual IFC objects and their conversion into DWH compatible elements are highlighted. Lastly it is also demonstrated that the combination of both technologies does not negatively reflect on the speed of database queries.

Keywords: BIM, IFC, Data Warehouse, Smart Metering, Key Performance Indicators

INTRODUCTION

Nowadays, buildings equipped with smart meters and sensors generate a huge amount of performance data. For evaluation purposes, this data needs to be processed and classified. The availability of standardized product and process models allows clustering of vast data by building elements and systems [1]. Sensor and meter readings enriched with spatial information increase their interpretability and meaningfulness [2].

In this research, we use information extracted from BIM models to classify performance data. Conventional BIM models require the analysis of complex 3D models. Our approach extracts only relevant information from an open, standardized BIM (IFC). For processing and complex analysis of IFC models, modern DWH functionality is implemented. The compatibility between an IFC model and DWH architecture allows an easy reuse of relevant BIM information. Classification of the data will then be possible through utilizing KPI's implemented in the DWH.

For this approach, only a subset of elements described in the IFC metadata model is required. The selected objects will be discussed and their successful transformation into a Data Warehouse is the main aspect of this work. Specifically, we select from the building model (i) spatial information, (ii) systems topology, (iii) relationship of different units within a company, and (iv) time reporting intervals. These four domains will be the foundation for the KPI's used for building classification.

BACKGROUND

IFC allows the export of metadata models into files. File types and layouts have been standardized. There is (i) IFC-STEP which has been standardized in ISO-10303-21. Every instance of an IFC object is stored in a single line. Each IFC object has a defined number of attributes. The STEP format enumerates these attributes chronologically, where unset attributes are represented with a dollar sign ($). Additionally, there is (ii) IFC-XML, a definition based on the widely used XML file type. The application of XML is the IFC domain is also standardized in ISO 10303-28. Due to the nature of the XML standard, the XML file comes with an overhead that significantly increases the total file size compared to STEP.

Currently, researchers are working on creating Data Warehouse servers which are compatible to IFC. Their vision is that IFC specifications can be considered as a blueprint for database schemas. Any useful information from an IFC file could then be easily imported into the schema. The expected outcome would be that information already modelled and stored in IFC could be reused for other applications. In the long term, this could eliminate repeated data acquisition for different purposes. The status and implementation of IFC compatible database schemas have been discussed in [3, 4].

Historical meter and sensor readings from buildings act as fact data in the proposed model. Preservation of this kind of data enables a wide field of disciplines. For example, energy analysts have data of much higher granularity for examination. Ways for data acquisition and data storage are outlined in previous research [5, 6].

METHODOLOGY

In this work, we aim to select a manageable subset of IFC objects extracted from the open BIM model to enrich the analysis of building performance data. These subsets can be clustered into four domains. For this, the IFC objects enumerated in table 1 will be of interest to us.

In IFC, a Model View Definition (MVD) is a specified subset of the IFC metadata model. MVD's allow the exchange of AEC information from selected individual domains. The principles of MVD's follow standards defined in ISO 29481. Unfortunately, for the model required by us, no MVD exists.

Domain	IFC Objects	IFC Relationship Objects
Spatial	IfcSpace, IfcBuildingStorey, IfcBuilding, IfcSite,	IfcRelAggregates
Organization	IfcOrganization,	IfcOrganizationRelationship
System	IfcDistributionFlowElement entities, IfcDistributionCircuit,	IfcRelAggregates
Time	IfcDateTimeResource entities, IfcTimeSeries	IfcResourceLevelRelationship

Table 1. Selected IFC objects

To compensate for this, we decided to focus solely on exported raw IFC data model files. This enables us to work with the bare data and select only individual IFC objects as needed. Both file type formats (STEP and XML) can be parsed using either on-board or 3rd party tools. In this work, we decided to use STEP files along with standard UNIX shell commands for parsing and filtering. Through these shell commands, all objects of the four domains as in table 1 can be filtered.

As an example, the objects consolidating the spatial domain are given in figure 1. There is illustrated how objects are linked to each other in IFC. Through resolving their hierarchical relationship, matching of a room to a storey and to a building becomes possible.

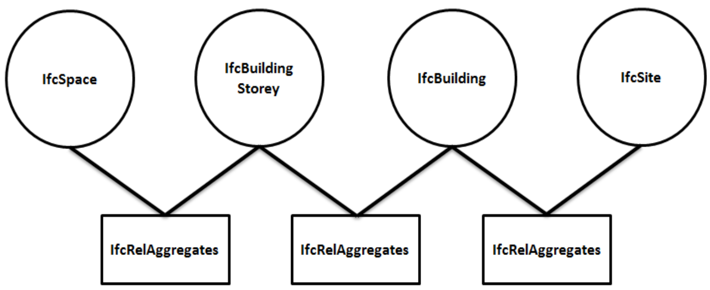

Figure 1. File based relation

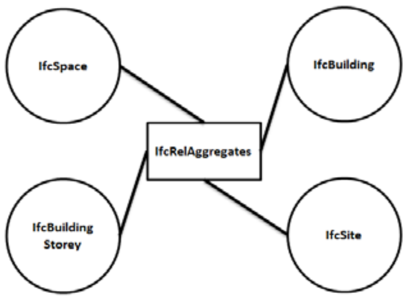

Figure 2. Database relation

To maintain compatibility between the IFC standard and the database, its schema is based upon the IFC object definition. IFC objects like IfcSpace and IfcBuilding hold their entities in dedicated tables. Relationship objects are consolidated in a single relationship table. These tables ensure that single objects, sitting in its own encapsulated tables, can be correctly related to their corresponding objects. Figure 2 highlights that in a database approach only one table IfcRelAggregates exists which maintains the relation of all related IFC entities.

Historical fact data is collected from meters and sensors throughout buildings and stored inside the DWH. IFC is supporting fact data through its IfcTimeSeriesValue entity. Each record consists of a value, a timestamp and meter/sensor identification (ID). Through these attributes, it is possible to identify which source reported what value at a specific time. The ID can be used to lookup further information from the geometrical model, e.g. where the sensor/meter is physically located.

The outlined remodelling process is executed for all four domains. At this stage, an IFC database schema compatible for our methodology has been achieved. However, this schema lacks speed due to its encapsulated design (as seen in the results section). In order to compensate this, we have implemented modern DWH technology on top.

The objects scattered across multiple tables from each domain have been consolidated in MV's (Materialized Views). This is accomplished by resolving the relationship table, effectively pulling all information into a single view. MVs get populated and updated in adjustable intervals and they allow the definition of constraints [7].

The fact table needs to be enriched with a) references to the MVs and b) with KPI's. This enrichment is also realised through the application of a MV. This work focuses on KPI's which were previously compiled and determined as beneficial for the analysis of building performance data [8, 9]. Figure 3 gives a graphical overview about earlier identified KPI categories.

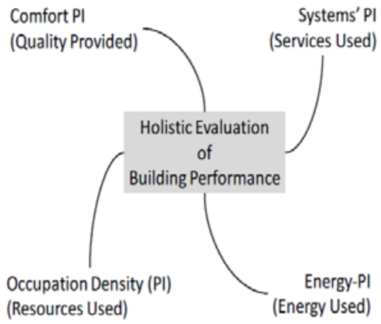

Figure 3. Categories of Key Performance Indicators as identified in [9]

Through the implementation of afore mentioned MV and the generation of a suitable fact table, all requirements are fulfilled. As a result of applying the constraints, a database star schema is created. For analysis purposes, OLAP cubes for each KPI can be generated based on this star schema. Example outputs (rendered in MATLAB) can be seen in the next section.

RESULTS

The selected results focus on the UPR (underperformance) KPI, which belongs to the class of comfort PI (as seen in figure 3). This KPI rates a building between 0% and 100%. In this context, 100% is equal with performance always within set thresholds. The two buildings evaluated are two university buildings, namely Civil and Environmental Engineering Building (CEE) and the Environmental Research Institute (ERI). Performance data for these buildings is collected since 2011. The following figures illustrate the exemplary analysis possibilities given by the outlined methodology.

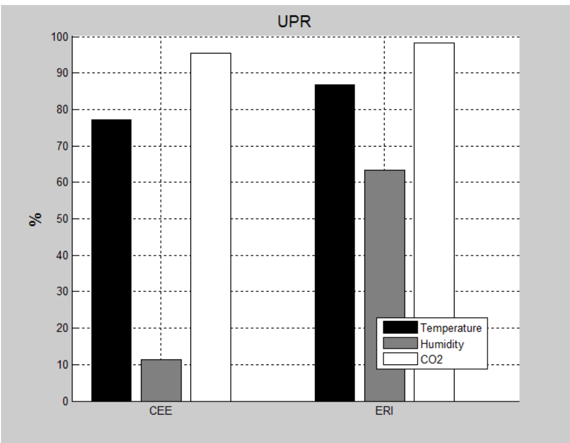

Figure 4a. Overall building comparison

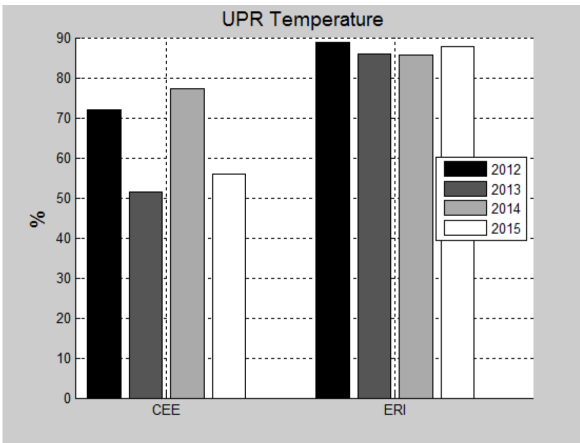

Figure 4b. Temperature comfort

The direct comparison between the two buildings reveals that the ERI buildings performance is superior in all three disciplines with the highest advance in humidity.

Graph 4b identifies clearly how the ERI building is consistently performing close to the desired levels of 100% comfort level, with only marginal changes of 2.5% over 4 years, while in contrast the level of user comfort in the CEE building is unpredictable with large variances in the data recorded

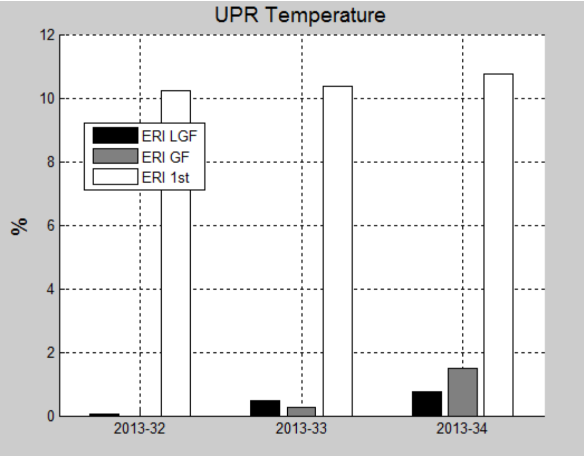

Figure 4c. Storey-based temperature

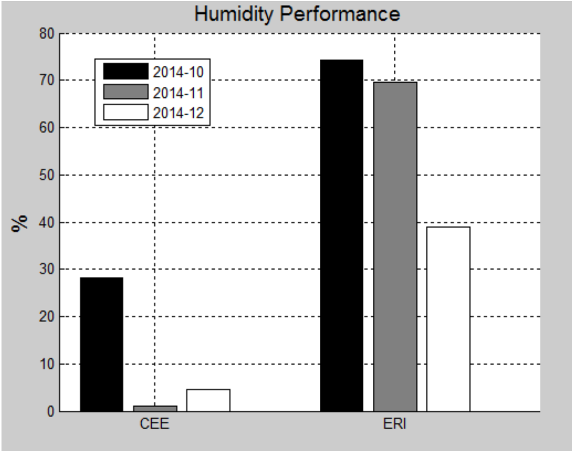

Figure 4d. Humidity comfort

Graph 4c is broken down into a 'story' spatial hierarchy. In this example, the amount of underperformance is displayed. This bad performance of the first floor may be due to its roof exposure.

When focusing on the performance of the two buildings in terms of humidity the CEE building is performing poorest with comfort far below the ERI in same time intervals.

Regular SQL queries to acquire KPI were executed against (i) a regular database schema without any DWH optimizations, and (ii) against a DWH optimized schema as introduced in this work. Both schemas held the same amount of fact data. Average response times can be seen in table 2.

Type of KPI	**Time span**	**Type of fact table**	
		Regular database (in s)	**DWH optimized (in s)**
Average consumption	1 month	75	0
Maximum consumption	1 day	6	0
Underperformance Ratio	1 day	8.7	0
Minimum value	6 month	83	0

Table 2. SQL execution comparison

The application of Data Warehouse technology is known to reduce response times to a minimum. These results prove that our methodology also benefits from this technology.

CONCLUSION

This work has introduced a methodology to process and classify building performance data using descriptive information from BIM models. This has been achieved by combining selected information from the open BIM standard IFC with DWH technology. This process eliminates the need to reacquire previously gathered information.

The introduced DWH dimensions allow the analysis of individual or grouped sensors/meters. This analysis is supported by the application of KPI. Analysis may be performed against arbitrary combinations of organization units, building systems, spatial zones over time periods. Even huge numbers of installed sensors can be analysed following the same principles.

A downside of this approach is, that even experienced database designers might face difficulties to fully understand the schema if they are not familiar with the IFC standard. Nonetheless, we believe that common schemas could improve rapid comparative analysis of collections of buildings.

For further application scenarios with different subsets of BIM data it is more than likely that other domains, as introduced in this paper, will be required. We are optimistic that the compatibility introduced in this work can be maintained. We envisage that a fully IFC compatible schema could improve efficiencies across the field. Database tools customized to work against IFC schemas could become a future niche product as they would work out-of-the-box with any stored data following the IFC meta-data model.

ACKNOWLEDGMENTS

This work is funded by the Irish HEA PRTLI-5 program and Bilfinger HSG FM.

REFERENCES

1. Ammar Ahmed, Joern Ploennigs, Karsten Menzel, Brian Cahill, Multi-dimensional building performance data management for continuous commissioning, Advanced Engineering Informatics, Volume 24, Issue 4, November 2010, Pages 466-475, ISSN 1474-0346, http://dx.doi.org/10.1016/j.aei.2010.06.007
2. Goekce, H.U., Goekce, K.U., 2013. Multi dimensional energy monitoring, analysis and optimization system for energy efficient building operations. Sustainable Cities and Society, Volume 10, February 2014, Pages 161-173, ISSN 2210-6707, http://dx.doi.org/10.1016/j.scs.2013.08.004
3. Mo, K., Menzel, K., and Hoerster, S. Development of an IFC-compatible data warehouse for a building information model. In: CiB W78. Beijing, 2013.
4. BaaS Deliverable 2.2 , Data Warehouse, EU FP7 project, 2013
5. Stack, P., Cahill, B., Manzoor, F. and Menzel, K., 2011.
Data Warehouse Model to support Optimized Operation and Energy Savings in Buildings. ICEBO, New York, 2011.
6. Hoerster, S., Katzemich, F., Menzel, K., 2014.
A Methodology for Data Logging and Retrieval from Remote Sites. ECPPM Vienna, 2014
7. Shukla, A., Deshpande, P., Naughton, J.F., Materialized view selection for multidimensional datasets, in: Proceedings of the 24th International Conference on Very Large Data Bases, NY, USA, 1998, pp. 488–499.
8. Menzel, K., Katzemich and F., Mahdavi, A. Impacts of building performance monitoring on integrated energy management. 2nd CESBP Vienna, 2013
9. Menzel, K., Browne, D. and Deng, S. Performance indicators to evaluate buildings' systems' performance. ECPPM Vienna, 2014